



## King's Research Portal

DOI:

[10.1016/j.procs.2011.12.012](https://doi.org/10.1016/j.procs.2011.12.012)

*Document Version*

Publisher's PDF, also known as Version of record

[Link to publication record in King's Research Portal](#)

*Citation for published version (APA):*

Lehrach, H., Subrak, R., Boyle, P., Pasterk, M., Zatloukal, K., Müller, H., Hubbard, T., Brand, A., Girolami, M., Jameson, D., Bruggeman, F. J., & Westerhoff, H. V. (2011). ITFoM - The IT future of medicine. In *Procedia Computer Science* (Vol. 7, pp. 26-29) <https://doi.org/10.1016/j.procs.2011.12.012>

### Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

### General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

### Take down policy

If you believe that this document breaches copyright please contact [librarypure@kcl.ac.uk](mailto:librarypure@kcl.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.

The European Future Technologies Conference and Exhibition 2011

**ITFoM – The IT Future of Medicine**

Hans Lehrach<sup>a,\*</sup>, Ralf Subrak<sup>a</sup>, Peter Boyle<sup>b</sup>, Markus Pasterk<sup>b</sup>, Kurt Zatloukal<sup>c</sup>,  
Heimo Müller<sup>c</sup>, Tim Hubbard<sup>d</sup>, Angela Brand<sup>e</sup>, Mark Girolami<sup>f</sup>, Daniel Jameson<sup>g</sup>,  
Frank J. Bruggeman<sup>h,i</sup>, Hans V. Westerhoff<sup>g,h</sup>

<sup>a</sup> Max Planck Institute for Molecular Genetics, Ihnestrasse 63-73, Berlin 14195, Germany<sup>b</sup> International Prevention Research Institute, Lyon, France<sup>c</sup> Medical University of Graz, Austria<sup>d</sup> Wellcome Trust Sanger Institute, Hinxton, UK<sup>e</sup> Institute for Public Health Genomics, Maastricht University, The Netherlands<sup>f</sup> Dept of Statistical Science, & Centre for Computational Statistics and Machine Learning, University College London, UK<sup>g</sup> Manchester Centre for Integrative Systems Biology, The University of Manchester, UK<sup>h</sup> The Netherlands' Institute for Systems Biology, VU University Amsterdam and the University of Amsterdam, NL<sup>i</sup> Life sciences, Centre for Mathematics and Computer Science, Amsterdam, NL

---

**Abstract**

Molecular medicine is undergoing a revolution, creating a data fog that may obscure understanding. The functioning human is analogous to a biological factory controlled by an incredibly complex Information and Communication (IC) network. It is proposed that 7 billion computational replicas be made of those 7 billion human IC networks to enable interrogation and manipulation, for understanding and personalized healthcare. This requires a revolutionary ICT that follows the organization of the biological information and communication flows, with implications for hardware, software and connectivity.

© conference organizers and published by Elsevier B.V. Open access under [CC BY-NC-ND license](#).

Selection and peer review under responsibility of FET11

**Keywords:** nonlinear information and communication technology; distributive computing; high throughput data analysis; personalized medicine; systems medicine; healthcare revolution: virtual human

---

**Introduction**

When forming foundations for large scale Information and Communication Technologies (ICT), computing science was driven by “large” physics and commercial applications, while medicine played a minor role. ICT requirements of the new, data-rich, individualized medicine will soon surpass the demands of all other (ICT) fields. As data-intensive analysis and computer-intensive modeling technologies become common clinical practice, medical ICT implementation will require more resources. The enormous demands of Monte Carlo simulations of biochemical systems and of the data-rich, individualized medicine poses unprecedented challenges to ICT, in hardware, storage, retrieval, communication, governance and policy.

The international research consortium ITFoM (Information and Communication Technology for Future Medicine) anticipates the medicine of the future, based on molecular, physiological and anatomical data from all individual

---

\* Corresponding author. Tel.: +49-30-8413-1221; fax: +49-30-8413-1380.

E-mail address: [lehrach@molgen.mpg.de](mailto:lehrach@molgen.mpg.de) (H. Lehrach).

patients. The 'ITFoM project' will help create the entirely new ICT that is necessary to enable models of human biochemical pathways, cells, tissues, diseases and ultimately of the human as a whole. These models correspond to the, Virtual Biochemical Human, from the molecular level upwards (molecules-up). ITFoM also associates closely with, and provides the missing biochemical basis for, the Virtual Physiological Human, which models at the physiological level (physiology down). Keys will be (i) the recognition that the essence of the physiological complexity of humans resides in their nature of being biological 'factories' organized by their own, complex IC'T' (where the 'T' is materially different from that in standard ICT and based on evolution rather than design), and (ii) the strategy of making ICT-based replicas of the molecular organization of individual humans. These individualized ICT-based replicas (from hereon ICT replicas) will then be used to identify personalized prevention/therapy schedules and side effects of drugs. ITFoM will prepare for the ICT-replica driven, amalgamation of three major areas: (i) medicine, from sample and diagnosis provision to clinical practice and patient consent, (ii) analytical techniques, covering functional genomics and imaging technology analyses on a routine basis, and (iii) ICT developments required to address the computational challenges.

The ICT replica is much more than a gigantic relational database of an individual. It is a dynamic real-time integration of a mathematical dynamic model simulating precisely the physiological functioning of the individual from molecules up with integrating all the molecular, genomic, medical and personal data of that individual. In real time it takes into account the common denominator of the replicas of all other humans on the planet, plus all available scientific knowledge about pathways, mechanisms, physics and chemistry, plus all track records of drug treatments.

### **Live-ICT based dynamic integration of dynamic biomedical data**

With the NIH (-USA) annual budget well over \$30 billion, how far are we with understanding the human, in terms of a precise 'virtual biochemical human' in terms of an ICT replica? We are not very far. The existing mathematical models are imprecise, ill connected, and nowhere near a molecule-based model for the whole body. The biomedical cottage industry does not significantly address the required integration of models and data. In DNA sequencing two paradoxical approaches put an end to a similar deadlock. One was the initiative to sequence the DNA of entire genomes linearly. The other was the shotgun approach. The genome sequencing initiative inspired a new way of thinking and thereby delivered much more quickly than anticipated. ITFoM tries to do the same for the deadlock of future molecular medicine: It recognizes that living organisms correspond to a complicated information and communication system comprising of huge numbers of computational units driving multiple factories. The hard coded information in the gene is conditionally communicated via mRNA to produce factory units (proteins). The computational logic is in the control elements (interactions) that dictate the level at which the gene is used to make proteins. The basic *human* program is made up of some 25,000 genes per cell. With  $1 \times 10^{14}$  cells in the human body it can be seen that we are potentially interested in the interactions between  $2.5 \times 10^{18}$  genes, extrapolating to some  $2.5 \times 10^{22}$  individual proteins. In order to understand the functioning of the human, a silicon replica of the human "IC system/computer" will be a key tool. With respect to healthcare in general this would be a 'virtual individual', with respect to therapy, a 'virtual patient'.

### **Managing the ICT challenge: Bio-inspired dynamically distributive ICT?**

Making an ICT replica of the human could be even more difficult than putting a man on the moon or sequencing the human genome. If all  $2.5 \times 10^{18}$  genes were to interact and each interaction required 10 calculations to evaluate, it would take all of today's supercomputers multiple lifetimes of the universe to compute one complete cycle of interactions for the human replica. Yet, the ICT challenge is to enable billions of replicas of the human to be evaluated for their life lifetime, i.e. 100 years, in a few seconds each. It may be possible to accelerate the production human ICT replica by implementing both software and hardware that follow the ICT architecture of the human itself. Metabolic pathways are full of calculations at the second time scale, whereas gene expression changes in tens of minutes. Biological computation and communication within a cell is much more intense than between cells. And, the  $2.5 \times 10^{18}$  genes are not all different, as each cell contains the same 25,000. Knowledge of the actual biology should come to the rescue of the otherwise impossible challenge of building all those ICT replicas. The strategy will be to build a 'common denominator replica' with an instantiation for each individual, ultimately 7 billion.

## **Meta ICT: integration of the required information and communication**

The ICT replicas should be realistic, and molecule based. A variety of analytical procedures are needed to produce massive amounts of data: the sequences of a human genome, the deep-sequenced transcriptome in multiple cells and tissues, images of the concentration distribution of proteins and ions in each individual. The ICT replica needs to integrate this with clinical information, e.g. diagnosis made by the physician, environmental, lifestyle and experience data of the patient, and population statistics. A third component is the various types of hardware, software, communication methodologies and information storage facilities: local storage arrays, bespoke server farms or generic cloud based information storage may be used at different dynamics. All these components and the human experts groups supplying them at high quality will exist at various locations on the planet. New dynamic networking technology will be needed to connect these flawlessly. Again the anatomy of the human may be followed here, attributing the various tissues or pathways to various research groups, with rapid computer communications between the web services of their servers.

## **Medical Dataflow: dynamic integration of a plethora of data types by the ICT replica**

The medical data formats considered in ITFoM comprises textual, narrative (diagnosis, environmental and lifestyle data, anamnesis, therapy), categorical (staging, grading, scores), hierarchical, network (family histories, pathway information), and numerical data (laboratory data, age, BMI), as well as signals (EEG, NMR, spectra), 2D images (histological images, X-Rays), 3D images (CT, MR, PET) and sequences/vectors (DNA sequences, PCR, molecular signatures). This poses major challenges in data compression, data exchange, data security, anonymization, ontology and clinical descriptions, especially for data related to disease phenotypes. For the development of the common denominator ICT replica of the human, ITFoM will provide a reference data set, tied in to internationally applicable criteria for data quality, in accordance with ethical and legal requirements. Depending on the clinical questions and the computational power, the individualized replicas will describe at various levels of abstraction, through a self-assembly process (sub models will ‘self-assemble’ from validated models and the reference data set).

ITFoM will consider the technological, medical, epidemiological, as well as the legal and governance aspects of collecting biological samples and medical data and unify existing (quality) standards and procedures, incorporating the work done by leading consortia and industry. The data sets for the different levels of the cellular control hierarchy (transcriptome, proteome, metabolome, etc.) will be integrated by the emerging ICT replica of the human. The medical dataflow will be organized in a ‘reversible mode’ such that medical scientists, clinicians as well as public health experts become motivated to initiate data collection and experiments that will further help develop the replica of the human: An interface of the developing human replica must be provided to this community for use in their daily practice of understanding health and disease. A dynamic “golden standard” dynamic data set that feeds into the developing human replica will be produced. The ICT replica of the individual will also contain mathematical models for the reconstruction of the data sets from a much more limited parameter set, enabling a drastic reduction of the data storage problem.

## **Interfaces**

The ITFoM reference data set will be continually extended as patients go through treatments and follow-up cycles, and will be made available, in anonymized, common denominator form to the public, and in personalized form to individuals. Both developments require specific procedures to deal with privacy, security, policy and governance issues and the development of a special informed consent. An ICT replica of each patient will become a key component of the health care system of the future. ICT replicas will be accessible through novel visualization and interaction technologies to all healthcare providers concerned with a particular patient and the patient himself (of course only with patient consent). The interface to an ICT replica will support different semantic views (expert, clinicians, patient, child, policymakers, etc.), multi language and visual interfaces, natural spoken interaction, multi-sense interface (touch the data) and the multi user capabilities (several experts able to discuss the same replica). A patient may not only be accompanied by the information on her/his genome, determined once during his lifetime, but also by the results of regular analysis of DNA/RNA, proteome and metabolome in body fluids. The monitoring and updating the ICT replica will need a specific interface. Collections of all ICT replicas that offer a window into population based analyses and provide a more rational basis for policy decisions, will require new tools for data discovery, mining and visualization. It will be challenging to support experts in the processes of surveillance and interpretation, hypothesis generation

and knowledge discovery of millions of ICT replicas. Pure visualization will fail to capture the vast amounts of heterogeneous data. Visual Analytics and new statistical methods will be required for data pre-processing, pre-filtering and clustering.

### **Challenges and perspectives**

Generation within 30 years of ICT replicas of the majority of individual humans on the planet would constitute the greatest ICT breakthrough ever, with huge ICT and medical spill-off. For the first time the enormous ICT implications of worldwide-individualized patient care will be addressed in combination with –omics technologies and medical requirements. The ICT replica will constitute a bonanza for the discovery of how the human molecular physiology produces function. And, the ICT replica of the human will lead to breakthroughs for medicine and healthcare as a whole. For the first time effects of possible therapies of disease can be predicted, new therapies can be designed, and therapies can be truly individualized, differing between individuals and between their diseases. Individualized medicine will become real, overtaking the cohort-based medicine that is developing at this moment.

Lower-hanging fruits are lurking: the individual ICT replica will put the heterogeneous data of millions of patients into a single context, i.e. their common denominator. By projection into the replicas, new statistics will have more power than today's non-structured statistics. The success of ICT replicas limited to single pathways, or diseases will motivate a spiral of further modeling, experimentation, assay development, patient analysis, and funding: From its early days, the ITFoM project will enable some calculation of health, disease, therapy for individual patients. This will begin to transform our health care with noticeable (i) benefits for health (prevention, diagnosis and therapy), (ii) reduction in cost by individualizing drug combinations, and (iii) new commercial opportunities in ICT, analytics and health care. The ultimate goal is: (i) To drive a revolution in ICT technologies by making them biomimetic. Relevant computing, storage, networking, and modeling technologies are developed to allow each patient's physician to have at minimum the power of analysis of a personal human genome project at diagnosis, treatment, and follow-up. (ii) A revolution in integrative information management and decision making that will allow the seamless connection of high throughput biomolecular characterization and clinical imaging technologies. The patient and her/his doctor in clinical applications, the drug researcher in discovery and development phases, the epidemiologist attempting to analyze health trends, and the public health experts and policymakers in developing effective national and EU wide health policies in a global context and with optimal governance, will all benefit. This entails the transformation of biomedical science from empirical and stochastic to fact based and knowledge driven i.e. based on an ICT paradigm that is in turn based on the marvelously effective yet complex organization of human life itself.

Supplementary material related to this article found, in the online version, at [doi:10.1016/j.procs.2011.12.012](https://doi.org/10.1016/j.procs.2011.12.012).

### **Acknowledgement**

The authors wish to thank the various research organizations that support the work that has led up to the ITFoM concept. IT Future of Medicine, (ITFoM) is funded by grant 285609, FP7.